

“MÉTODO FL-PMC (FOURIER LAPPED – PERCEPTRON MULTICAMADAS) PARA A ESTIMAÇÃO DE QUALIDADE DE VOZ”

A presente patente de invenção consiste de um método
5 objetivo para a estimação confiável da qualidade subjetiva de sinais de voz na faixa de telefonia. Este método apresenta melhor desempenho que os métodos tradicionais e menor custo computacional.

Os métodos objetivos para a estimação da qualidade de voz são empregados para a otimização, homologação ou monitoração de
10 codecs (codificadores e decodificadores) de voz e outros equipamentos e sistemas de telecomunicações. Tais tarefas eram tradicionalmente realizadas através de testes subjetivos, complexos, caros e que demandam infraestrutura laboratorial especializada. Isto motivou a pesquisa internacional por procedimentos objetivos que apresentassem resultados
15 altamente correlacionados com as medidas subjetivas.

Os métodos objetivos de avaliação da qualidade de voz disponíveis, quando comparados aos métodos subjetivos, apresentam as seguintes vantagens: (a) custo baixo, (b) fácil utilização e (c) tempo de execução praticamente desprezível em aplicações *on-line* ou *off-line*.
20 Baseadas nessas vantagens, diferentes abordagens foram propostas. As principais são destacadas a seguir:

- PSQM [REF2] (Perceptual Speech Quality Measure): o primeiro padrão adotado pela ITU (International Telecommunication Union), amplamente utilizado;
- 25 • PSQM+ [REF7] (Perceptual Speech Quality Measure PLUS): uma versão melhorada do método PSQM;
- PAMS [REF3] (Perceptual Analysis Measurement System): primeiro método a considerar os atrasos variáveis entre os sinais

de voz original e degradado;

- PESQ [REF4] (Perceptual Evaluation of Speech Quality): avalia a degradação na transmissão de sinais de voz diante de condições críticas das redes de telecomunicações (Voz sobre IP, ISDN, xDSL e GSM).

O método FL-PMC objetivo, objeto da presente patente, obtém melhores níveis de correlação entre suas estimativas de qualidade e a qualidade subjetiva a ser estimada, que os métodos existentes. Além, da melhora de desempenho, o método FL-PMC, objeto da presente patente, é executado com baixo esforço computacional. Estas vantagens foram obtidas a partir de duas contribuições inovadoras: (a) desenvolvimento de novas formas de extração de parâmetros objetivos dos sinais de voz; e (b) aplicação de uma solução não-linear para o problema de mapeamento multidimensional entre os parâmetros extraídos dos sinais de voz e o valor numérico correspondente à qualidade subjetiva.

A figura 01 anexada mostra a disposição em diagrama de blocos do método FL-PMC, objeto da presente patente. A seguir é descrita a funcionalidade de cada bloco mostrado na figura 01.

- SINAL DE VOZ ORIGINAL: representa o sinal de voz a ser transmitido;
- SINAL DE VOZ DEGRADADO: representa o sinal de voz original após ter sido transmitido e submetido a diferentes tipos de agentes degradantes, tais como: codecs de voz, perdas de quadros, ruído de ambiente, condições críticas de funcionamento de canais de comunicação móvel, etc;
- PRÉ-PROCESSAMENTO: realiza algumas tarefas de processamento nos sinais de voz original e degradado, como o alinhamento temporal entre os sinais (1) e o escalonamento para igualar

suas energias (2);

- FFT (FAST FOURIER TRANSFORM) (3):

aplicação da transformada rápida de Fourier aos sinais de voz e adaptação dos padrões resultantes às características auditivas humanas;

- 5 • MLT (MODULATED LAPPED TRANSFORM)

(4): aplicação da transformada Lapped Modulada [REF5] aos sinais de voz e adaptação dos padrões resultantes às características auditivas humanas;

- EXTRAÇÃO DE PARÂMETROS: cinco diferentes medidas são extraídas a partir dos padrões resultantes da aplicação das transformadas e de sua adaptação às características da audição humana, resultando num total de dez parâmetros. Os parâmetros extraídos são: a diferença entre as SNR (*signal to noise ratio*- relação sinal ruído) segmentadas entre os sinais original e degradado (5), a distância espectral perceptual [REF6] (6), a distância cepstral perceptual [REF6] (7), a medida objetiva MOQV1 [REF1] (8) e a medida objetiva MOQV2 [REF1] (9);

- MAPEAMENTO NÃO-LINEAR (10): implementação da rede neural Perceptron multicamadas [REF8] para realização do mapeamento multidimensional e não-linear entre os dez parâmetros extraídos dos sinais de voz e o valor correspondente de qualidade subjetiva. A rede neural Perceptron multicamadas é implementada com 10 entradas, correspondentes aos dez parâmetros extraídos dos sinais de voz, uma camada oculta com um número de neurônios entre 10 e 50 e, finalmente, uma camada de saída com as seguintes configurações de saída:

- 25 • Saída 1: a rede neural Perceptron multicamadas possui apenas uma saída. Durante o processo de aprendizagem a rede neural foi treinada para que sua única saída estime um valor numérico e real correspondente ao valor MOS [REF1] ou CMOS [REF1];

- Saída 2: a rede neural Perceptron multicamadas possui quatro saídas. Para esta configuração de saída, a rede neural foi treinada para que as suas quatro saídas estimem um código binário de quatro bits. O código binário de quatro bits corresponde a um dos 16 níveis distintos, possíveis e obtidos com a combinação dos quatro bits. Além disso, os 16 níveis distintos representam a quantização da faixa de valores MOS ou CMOS em 16 níveis discretos. A estimativa do valor MOS ou CMOS é obtida a partir da relação do código binário estimado pela rede neural e o seu correspondente nível discreto;
- 10 • Saída 3: a rede neural Perceptron multicamadas possui cinco saídas. A descrição é a mesma dada para a configuração Saída 2, com a ressalva de que as cinco saídas da rede estimam um código binário de cinco bits;
- Saída 4: a rede neural Perceptron multicamadas possui seis saídas. A descrição é a mesma dada para a configuração Saída 2, com a ressalva de que as seis saídas da rede estimam um código binário de seis bits;
- 15 • Saída 5: a rede neural Perceptron multicamadas possui sete saídas. A descrição é a mesma dada para a configuração Saída 2, com a ressalva de que as sete saídas da rede estimam um código binário de sete bits;
- Saída 6: a rede neural Perceptron multicamadas possui oito saídas. A descrição é a mesma dada para a configuração Saída 2, com a ressalva de que as oito saídas da rede estimam um código binário de oito bits;
- 20

- QUALIDADE SUBJETIVA ESTIMADA:

Armazena o valor numérico correspondente à estimativa da qualidade subjetiva dos sinais de voz oferecida pelo mapeamento implementado pela rede neural Perceptron multicamadas.

REFERÊNCIAS BIBLIOGRÁFICAS:

[REF1] J. G. A. Barbedo. "Objective quality assessment of telephone-band speech codecs (in Portuguese)," *Master's Thesis, Unicamp, Campinas*,

July, 2001.

[REF2] J. G. Beerends, J. A. Stemerdink, "A perceptual speech-quality measure based on a psycho acoustic sound representation," *Journal of the Audio Engineering Society*, 42(3), March, 1994, pages. 115-123.

5 [REF3] PAMS website: <http://www.labs.bt.com/people/rixa/pams/>.

[REF4] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, 2001.

[REF5] H. S. Malvar, *Signal Processing with Lapped Transforms*, Artech
10 House, MA: Norwood, 1992.

[REF6] A.V. Oppenheim, R.W. Schaffer, *Discrete Time Signal Processing*, Prentice Hall, New Jersey, 1989.

[REF7] International Telecommunication Union website:
<http://www.itu.int/>.

15 [REF8] S. Haykin, *Neural Networks A Comprehensive Foundation*. Prentice Hall, New Jersey, 1999.

REIVINDICAÇÕES

- 1) “MÉTODO FL-PMC (FOURIER LAPPED – PERCEPTRON MULTICAMADAS) PARA A ESTIMAÇÃO DE QUALIDADE DE VOZ” caracterizado por apresentar as seguintes técnicas: alinhamento temporal dos sinais de voz (1); escalonamento global para igualar a energia dos sinais de voz (2); aplicação da transformada rápida de Fourier (FFT) nos sinais de voz (3); aplicação da transformada Lapped Modulada (MLT) nos sinais de voz (4); extração da diferença entre as SNR (*signal to noise ratio* - relação sinal ruído) segmentadas dos sinais de voz (5); extração da distância espectral perceptual (6), a partir dos dados obtidos após a aplicação da FFT e da MLT; extração da distância cepstral perceptual (7), a partir dos dados obtidos após a aplicação da FFT e da MLT; cálculo da medida objetiva MOQV1 (8), a partir dos dados obtidos após a aplicação da FFT e da MLT; cálculo da medida objetiva MOQV2 (9), a partir dos dados obtidos após a aplicação da FFT e da MLT; implementação da rede neural Perceptron multicamadas para estimar um valor numérico correspondente à qualidade subjetiva (MOS – *Mean opinion score* ou CMOS – *Comparative mean opinion score*) dos sinais de voz (10).
- 2) “MÉTODO FL-PMC (FOURIER LAPPED – PERCEPTRON MULTICAMADAS) PARA A ESTIMAÇÃO DE QUALIDADE DE VOZ” caracterizado por usar técnicas de processamento de sinais implementadas, primeiramente, para o pré-processamento e adequação dos sinais de voz original e degradado, (1) e (2), seguido da aplicação das técnicas FFT (3) e MLT (4) nos sinais de voz e da adaptação dos padrões resultantes às características auditivas humanas, a fim de se extrair diferentes parâmetros dos sinais de voz.
- 3) “MÉTODO FL-PMC (FOURIER LAPPED – PERCEPTRON MULTI-

- CAMADAS) PARA A ESTIMAÇÃO DE QUALIDADE DE VOZ” caracterizado pelo uso de técnicas de processamento de sinais empregadas para o cálculo da diferença da SNR segmentada (5) entre os sinais original e degradado, da distância espectral perceptual (6), da distância cepstral perceptual (7), do valor objetivo MOQV1 (8) e do valor objetivo MOQV2 (9). Esses valores calculados, após a aplicação da FFT e MLT, representam as informações ou parâmetros qualitativos e quantitativos extraídos dos sinais de voz, a fim de serem utilizados como variáveis de entrada na técnica de mapeamento.
- 4) “MÉTODO FL-PMC (FOURIER LAPPED – PERCEPTRON MULTICAMADAS) PARA A ESTIMAÇÃO DE QUALIDADE DE VOZ” caracterizado pelo uso da rede neural Perceptron multicamadas (10) para estimar o valor numérico (MOS – *Mean opinion score* ou CMOS – *Comparative mean opinion score*) correspondente à qualidade subjetiva dos sinais de voz através de um mapeamento multidimensional e não-linear entre os parâmetros extraídos dos sinais de voz e a correspondente medida de qualidade subjetiva dos sinais de voz. Essa rede neural possui 10 entradas, as quais representam os dez parâmetros extraídos do sinal, uma única camada intermediária, com número de neurônios entre 10 e 50 e uma camada de saída com diferentes configurações.

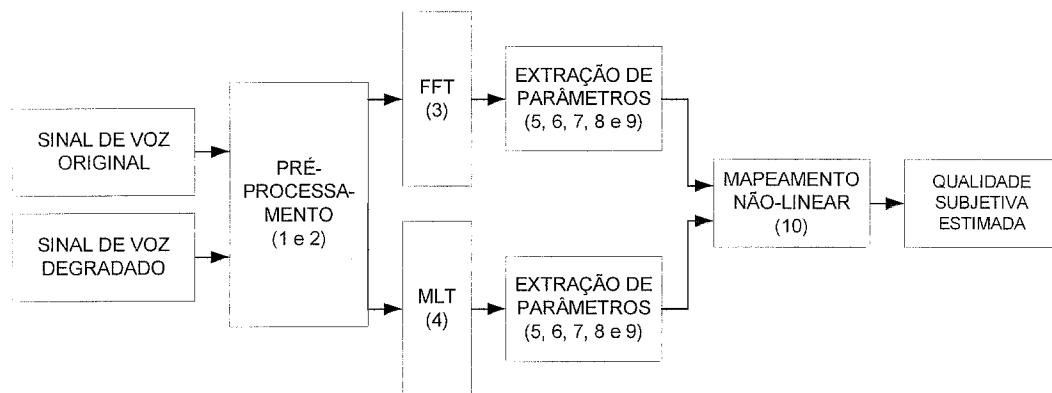


FIG. 01

RESUMO

“MÉTODO FL-PMC (FOURIER LAPPED – PERCEPTRON MULTICAMADAS) PARA A ESTIMAÇÃO DE QUALIDADE DE VOZ”. A Patente de invenção para a estimação da qualidade subjetiva dos

5 sinais de voz é compreendida por: a) técnicas de processamento de sinais aplicadas ao alinhamento (1) e ajustes para igualar o nível de energia (2) dos sinais de voz; aplicação das transformadas FFT (3) e MLT (4) e adaptação dos padrões resultantes às características auditivas humanas, a fim de revelar as informações espectrais e perceptuais dos sinais de voz;

10 extração da diferença entre as SNR segmentadas (5) entre os sinais de voz; extração da distância espectral perceptual (6), após a aplicação da FFT e da MLT; extração da distância cepstral perceptual (7), após a aplicação da FFT e da MLT; cálculo da medida objetiva MOQV1 (8) após a aplicação da FFT e da MLT e cálculo da medida objetiva MOQV2 (9) após a

15 aplicação da FFT e da MLT; implementação da rede neural Perceptron multicamadas para o mapeamento multidimensional e não-linear dos dez parâmetros extraídos dos sinais de voz numa correspondente estimativa da qualidade subjetiva dos sinais de voz (MOS ou CMOS) ou por um código binário associado aos valores MOS e CMOS (10).